

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

1 9 9 9 年 2 月 4 日

出 願 番 号
Application Number:

平成 1 1 年 特 許 願 第 0 2 7 7 3 6 号

出 願 人
Applicant (s):

オリンパス光学工業株式会社



2 0 0 0 年 9 月 2 2 日

特 許 庁 長 官
Commissioner,
Patent Office

及 川 耕 造



出 証 番 号 出 証 特 2 0 0 0 - 3 0 7 6 0 8 6

【書類名】	特許願
【整理番号】	99P00270
【提出日】	平成11年 2月 4日
【あて先】	特許庁長官 殿
【国際特許分類】	C12Q1/68 C12N15/09 G01N33/50
【発明の名称】	遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法
【請求項の数】	15
【発明者】	
【住所又は居所】	東京都八王子市下柚木 3 - 6 - 5 0 1
【氏名】	陶山 明
【特許出願人】	
【識別番号】	000000376
【氏名又は名称】	オリンパス光学工業株式会社
【代表者】	岸本 正壽
【手数料の表示】	
【予納台帳番号】	002314
【納付金額】	21,000円
【提出物件の目録】	
【物件名】	明細書 1
【物件名】	図面 1
【物件名】	要約書 1
【プルーフの要否】	要

【書類名】 明細書

【発明の名称】 遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法

【特許請求の範囲】

【請求項 1】

繰り返し塩基配列を生ずるような少ない配列数を 1 単位とし、既知の核酸分子を 1 単位毎にもれなく分割し、分割した個々の単位配列の出現頻度に基づいて、遺伝子情報解析に適した単位配列を少なくとも含むような塩基配列を解析用配列として抽出することを特徴とする遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法。

【請求項 2】

前記抽出する工程を、複数の異なるアルゴリズムを順次適用して実行することを特徴とする請求項 1 に記載の方法。

【請求項 3】

特定の配列数を 1 単位として抽出することで 1 サイクルとし、サイクル毎に配列数が多くなるように設定した単位配列によって、複数サイクルの抽出工程を実行することを特徴とする請求項 1 に記載の方法。

【請求項 4】

抽出工程が、さらに単位配列毎の化学的特性に基づいて取捨選択する工程を有することを特徴とする請求項 1 に記載の方法。

【請求項 5】

化学的特性が分子学的安定性であることを特徴とする請求項 4 に記載の方法。

【請求項 6】

化学的特性が単位配列の熱的安定性であることを特徴とする請求項 5 に記載の方法。

【請求項 7】

化学的特性が単位配列の T_m 値および／または自己の二次構造安定性であることを特徴とする請求項 5 に記載の方法。

【請求項 8】

n 個 (n は 2 以上の整数) の塩基配列からなる n 単位配列の全ての組合せに相

当する 4^n 種類の中で、既知の核酸分子の塩基配列上に出現する各 n 単位配列の出現頻度を演算する第 1 の演算工程と、

前記第 1 演算工程で得た出現頻度のうち少なくとも最も出現頻度が低い n 単位配列を抽出する第 1 の抽出工程と、

抽出された低頻度・単位配列を含むように m 個 (m は 1 以上の整数) 多い p 個の塩基配列からなる p 単位配列の全ての $4^m \times p$ 種類の中で、前記核酸分子に出現する各 p 単位配列の出現頻度を演算する第 2 の演算工程と、

前記第 2 演算工程で得た出現頻度が低いような単位配列を解析用配列として抽出する第 2 の抽出工程とを有することを特徴とする遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法。

【請求項 9】

第 2 演算工程が、この第 2 演算工程で得た最も出現頻度の低い p 単位配列を少なくとも選択し、この選択した p 単位配列を含むような、より多い個数の単位配列に関して同様の演算を少なくとも 1 回繰り返す繰り返し工程をさらに含むことを特徴とする請求項 8 に記載の方法。

【請求項 10】

n は、4、5 または 6 である請求項 8 に記載の方法。

【請求項 11】

m は、2 以上である請求項 8 に記載の方法。

【請求項 12】

少なくとも第 2 の抽出工程が、複数の低頻度・ p 単位配列の中から化学的条件に基づいて取捨選択する工程をさらに有することを特徴とする請求項 8 に記載の方法。

【請求項 13】

化学的条件が T_m 値の安定性および／または二次構造の安定性であることを特徴とする請求項 12 に記載の方法。

【請求項 14】

取捨選択する工程が、所定範囲の T_m 値に含まれる p 単位配列を複数個選択し、次いで、 T_m 値により選ばれた複数の p 単位配列の中から 2 次構造が安定なも

のを選択する工程を含むことを特徴とする請求項 12 に記載の方法

【請求項 15】

請求項 1 または 8 の決定方法における全ての工程を、コンピュータにより順次実行することを特徴とする方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、特定の遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法に関する、特に、膨大な遺伝子情報の中から短時間で効率良く決定する方法に関する。本発明は、限られた計算能力で演算できるようなアルゴリズムを複数組合せて、目的の塩基配列を簡便かつ効率良く決定できるので、遺伝子情報解析に極めて有効に利用できるものである。

【0002】

【従来の技術】

分子生物学において、生体試料中に存在する特定核酸分子の検出は、特定臓器で発現・機能する蛋白質の核酸分子レベルでの解析、神経、脳あるいは免疫系での情報伝達での蛋白質の発現制御の研究等において、重要であるのみならず、遺伝子病の変異遺伝子検出、癌の診断、ウイルス関連遺伝子の検出など遺伝子診断において重要な技術である。特に、遺伝子診断は、確定診断として用いられるため誤りが許されない。

【0003】

種々の核酸分子を定性的に又は定量的に検出することによって遺伝子の存在や変異等を解析する遺伝子解析技術にとって、標的核酸と特異的に結合するような塩基配列を有する核酸（プローブ）配列（以下、解析用配列と称する）を決定することは最も重要な事項である。この解析用配列への要望は、核酸ハイブリダイゼーション反応を応用した解析技術の進歩に伴って益々強くなってきている。核酸ハイブリダイゼーション反応は、標的核酸の異なる位置で類似する塩基配列を有するような解析用配列を用いると、ミス・ハイブリダイズし易くなる。また、自己の二次構造の安定性が高い解析用配列を用いると、ハイブリダイゼーション

の効率ないし正確性が低下する。従って、解析用配列には、なるべくユニークな塩基配列となるように、全長を長く設定する方法が採られがちである。しかし、解析用配列は、全長が長くなればなるほど、それ自身の二次構造がより安定化した特異的形状を持つようになって、ハイブリダイズ不可能となるか或いはハイブリダイゼーション温度が高くなり過ぎて実験や検査等が困難になる。さらに、これら多くの分子生物学的技術において用いられる解析用配列の選択には、多くの経験や試行錯誤を要し、かつ従来の解析用配列の決定に用いる計算方法では、適切な核酸配列の選択に膨大に長い計算時間を要するものであった。

【0004】

【発明が解決しようとする課題】

本発明は、常に正確なハイブリダイズ反応を示すような解析用配列を決定する方法を提供することを目的とする。また、本発明は、解析用配列を迅速に決定できる方法を提供することを目的とする。また、本発明は、大型コンピュータを用いることなく、簡便にかつ経済的に有利な解析用配列の決定方法を提供することを目的とする。

【0005】

【課題を解決するための手段および作用効果】

上述した課題を解決するために、本発明は、繰り返し塩基配列を生ずるような少ない配列数を1単位とし、既知の核酸分子を1単位毎にもれなく分割し、分割した個々の単位配列の出現頻度に基づいて、遺伝子情報解析に適した単位配列を少なくとも含むような塩基配列を解析用配列として抽出することを特徴とする遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法である。

【0006】

ここで、抽出する工程は、複数の異なるアルゴリズムを順次適用して実行するのが好ましい。また、特定の配列数を1単位として抽出することで1サイクルとし、サイクル毎に配列数が多くなるように設定した単位配列によって、複数サイクルの抽出工程を実行するようにしてもよい。

【0007】

また、抽出工程は、さらに単位配列毎の化学的特性に基づいて取捨選択する工

程を有していてもよい。この場合には、解析用塩基配列としての反応性の適否に応じた選択ができるので、有効な配列決定ができる。特に、化学的特性が分子学的安定性であるようにすれば、ハイブリダイゼーション反応の適否に基づいた選択が可能となる。化学的特性としては、単位配列の熱的安定性および二次構造の安定性のいずれか又は両方であるのが好ましい。

【0008】

また、本発明は、 n 個（ n は2以上の整数）の塩基配列からなる n 単位配列の全ての組合せに相当する 4^n 種類の中で、既知の核酸分子の塩基配列上に出現する各 n 単位配列の出現頻度を演算する第1の演算工程と、前記第1演算工程で得た出現頻度のうち少なくとも最も出現頻度が低い n 単位配列を抽出する第1の抽出工程と、抽出された低頻度・単位配列を含むように m 個（ m は1以上の整数）多い p 個の塩基配列からなる p 単位配列の全ての $4^m \times p$ 種類の中で、前記核酸分子に出現する各 p 単位配列の出現頻度を演算する第2の演算工程と、前記第2演算工程で得た出現頻度が低いような単位配列を解析用配列として抽出する第2の抽出工程とを有することを特徴とする遺伝子核酸配列の検出に用いる核酸塩基配列の決定方法である。

【0009】

ここで、第2演算工程が、この第2演算工程で得た最も出現頻度の低い p 単位配列を少なくとも選択し、この選択した p 単位配列を含むような、より多い個数の単位配列に関して同様の演算を少なくとも1回繰り返す繰り返し工程をさらに含むようにすれば、配列決定の効率がより向上するので好ましい。また、 n は、4、5または6のいずれかであれば、出現頻度を求める塩基配列の組合せが、それぞれ256通り（ $n=4$ ）、1024通り（ $n=5$ ）または4096通り（ $n=6$ ）となって、演算数を実用的な処理速度で実行可能となるので好ましい。また、 m は、2以上であれば、第1の抽出工程で選定された n 単位配列を含むユニークな塩基配列を効率良く探す上で有利である。 p 単位配列は、 n 単位配列を含むように既知の標的核酸中に実際に存在する部分配列として選ばれるので、比較的長い配列、例えば、 $m=10 \sim 50$ の範囲に設定することができる。

【0010】

また、少なくとも第2の抽出工程は、複数の低頻度・p単位配列の中から化学的条件に基づいて取捨選択するのが好ましい。ここで、化学的条件は、分子学的安定性を判定基準にするのが好ましく、具体的には、T_m値の安定性および／または自己の二次構造の安定性であるのが好ましい。特に、T_m値と二次構造の両方の安定性で選択する場合には、取捨選択する工程が、所定範囲のT_m値に含まれるp単位配列を複数個選択し、次いで、T_m値により選ばれた複数のp単位配列の中から2次構造が安定なものを選択する工程を含むようにするのが好ましい。

【0011】

さらに、上述した決定方法は、比較的少ない演算処理量によって実行できるので、全ての工程を、コンピュータにより順次実行するようにして、簡単かつ安価に解析用配列を決定することができる。ここで、大型コンピュータのような大容量の計算速度を要せず、一般に普及しているような中型（デスクトップ型）、ひいては小型（ノート型）のコンピュータで実行できる利点がある。

【0012】

一般に、相同性が有る配列は、相補的結合を示すが、ハイブリダイズするためには、5個以上の塩基配列が連続して一致しなければ安定な結合を示さない。換言すると、相補的配列とは、例えば6塩基からなる核酸配列が、6箇所中5箇所でハイブリダイズし、1箇所においてのみ塩基配列が一致しない場合をいう。従って、5塩基が一致する場合に、ミス・ハイブリダイゼーションが発生し得る。従来の解析用配列の設計は、このようなミス・ハイブリダイゼーションを十分に考慮していない。既知の標的核酸に対して設計される従来の解析用配列は、多かれ少なかれ相補的配列を有していながらも、ミス・ハイブリダイゼーションの可能性を有している。これに対して、本発明は、ミス・ハイブリダイゼーションの仕組みを考慮した解析用配列を決定する点を重要な目的としている。

【0013】

二次構造安定性とは、少なくともハイブリダイゼーション反応において、解析用配列が立体構造を形成せずに、正しいハイブリッドを形成する性質をいう。解析用配列が立体的構造を取って自己安定化すると、所望のハイブリッドを形成不

可能になる。ここで、立体的構造とは、核酸鎖によるループ化、核酸鎖同士の交差結合化等をいう。かかる立体的構造を取らない核酸鎖は、遺伝子情報解析過程において、標的遺伝子にのみ特異的かつ効率良く結合する。

【0014】

【発明の実施の形態】

以下に、本発明の方法を実施の形態に沿って説明するが、本発明は、以下の説明に限定されるものではない。

【0015】

本発明の実施の形態を図1を用いて説明する。図1は、本発明の方法を簡単かつ効率良く実施するために、解析用配列としての遺伝子同定プローブのプローブ配列を設計する場合を例に、模式的に示したフローチャートである。特に、この実施の形態は、計算量がゲノムの長さに比例する高速アルゴリズムを提供するものである。

【0016】

まず、既知の標的核酸分子のORF (open reading frame) 領域（以下、図1およびその説明では単に遺伝子と称する）に表現されている塩基配列について、5個の塩基配列からなる単位配列（以下、5・タプルと称する）の全ての組合せを得る。このような全ての5・タプルの組み合わせは、図2に示すように、ゲノム1の所定の遺伝子領域端部から5塩基分の塩基配列を個別のタプル2として、1塩基づつずらして得られる。ここで、遺伝子の塩基配列を構成する4種類の塩基（A、T、C、G）は、任意の遺伝子に関し、最大でも $4^5 (= 1024)$ 通り（即ち、単位配列を構成する塩基数 n について最大 4^n 通り）の組合せである。従って、ヒトゲノムのように高度に複雑な遺伝子を有する場合でも、高々1024通りの組合せが得られる。大腸菌のように、複雑性が低いゲノムでは、1024通りに満たない組合せが得られる。

【0017】

次に、得られた全ての組合せの5・タプルについて、塩基配列毎に分別して遺伝子上での出現頻度を統計計算する（第1の演算工程）。この場合、図4に示すように、縦軸に出現頻度を、横軸に各種塩基配列に分別したタプルを当てたグラ

フを作成すれば、比較検討し易い。本発明では、ミス・ハイブリダイゼーションを回避するために、出現頻度が低い塩基配列に注目し、少なくとも最も出現頻度が低い塩基配列からなる 5・タプル（以下、低頻度タプルと称する）を抽出する（第 1 の抽出工程）。出現頻度の最下位が複数有る場合には、これら複数の低頻度タプルを抽出するのが好ましい。なお、このような計算とグラフ化および抽出作業は、市販のコンピュータで容易に実行可能である。また、5・タプルの各塩基配列毎の出現頻度は、メモリ等に記憶しておくものとする。

【0018】

次に、5・タプルよりも 2 個以上、好ましくは 10～50 個分多い塩基、例えば 30 塩基からなる単位配列であって、かつ低頻度タプルが発現している位置を含むようなゲノム 1 の所定遺伝子領域上の配列（以下、部分配列と称する）を、図 3 に示すように低頻度タプルの前後双方にずらすことによって、多数の個別の部分配列 3 を得る。ここで、5・タプルを含む 30 塩基の全ての組合せは、最大 $4^{25} \times 30$ 通り（即ち、タプルを構成する塩基数 n よりも m 個多い、合計 p 個の塩基数からなる単位配列である部分配列について、最大 $4^m \times p$ 通り）考えられるが、低頻度タプルを含むような部分配列の前後 25 塩基（即ち、前後 m 塩基）についてずらすだけなので、遺伝子上の組合せは非常に少ない組合せとなるので、所要計算時間が顕著に短縮される。このような遺伝子上の部分配列は、図 4 で説明したのと同様にして、グラフ化し、塩基配列毎の出現頻度を計算する（第 2 の演算工程）。本発明では、ミス・ハイブリダイゼーションを回避するために、出現頻度が低い塩基配列に注目し、最も出現頻度が低い塩基配列を含んで、最下位から順に複数の部分配列（以下、低頻度部分配列と称する）を抽出する（第 2 の抽出工程）。出現頻度の最下位が複数有る場合には、これら複数の低頻度部分配列を抽出するのが好ましい。なお、このような計算とグラフ化および抽出作業は、市販のコンピュータで容易に実行可能である。また、部分配列の各塩基配列毎の出現頻度は、メモリ等に記憶しておくものとする。

【0019】

次に、抽出された低頻度部分配列群について、出現頻度以外の条件、即ち物理化学的条件によって、目的のプローブ配列を取捨選択する。このように、プロー

ブ配列を単に出現頻度のみで決定しない理由は、必ずしも、最もユニークな塩基配列を有するプローブ配列が強固にかつ効率良くハイブリッドを形成するとは限らないからである。そこで、具体的には、図5および図6に示すように、複数の低頻度部分配列について、それぞれの熱的安定性を検討するのが好ましい。まず、図5のように、 T_m 値をグラフ化して、所定の T_m 値範囲に収まるような変動の少ない低頻度部分配列を取捨選択する。ここで、塩基配列によっては、 T_m 値がばらつき、これに起因してハイブリダイゼーションの熱的変化の過程で結合が外れ易くなるので、このような T_m 値がばらつく低頻度部分配列を除去する。 T_m 値の計算は、塩基配列から算定することができる。次に、取捨選択により除去されなかった残りの低頻度部分配列を有力候補として、ハイブリダイゼーション反応の熱的変化に対する二次構造の安定性を検討する。

【0020】

ここで、塩基配列によっては、図6に示すように、適宜のハイブリダイゼーション環境、例えば、固相としての支持体4上に適当なリンカー5を介して候補配列を固定して被検サンプル等の反応成分を含んだ混合液中に置かれた条件の下で、上記ループ状の形態に変形する部分配列6や、同一の低頻度部分配列同士で部分的な相補結合を生じる部分配列7が、それ以上ハイブリダイズしないか或いは誤ったハイブリダイゼーションを生じるような別形態になって安定化してしまう場合がある。従って、このような安定化二次構造を形成し得るものを除去して、ハイブリダイゼーション環境下でも、安定化二次構造に変形または部分的結合を生じずに、単独の状態を維持するような部分配列8を最有力候補として選択する。なお、二次構造の安定性は、適当な解析用ソフトウェアを用いて、塩基配列から算定することもできる。

【0021】

最後に、上述した出現頻度および物理化学的条件により絞り込まれた特定の低頻度部分配列を、遺伝子同定用のプローブ配列として使用可能かどうか確認する。この確認工程は、遺伝子の全長を用いて、選択したプローブ配列が唯一の特定位置に相補的に結合するか否かによって実行することができる。かかる結合特異性もまた、遺伝子全体の塩基配列と対比するようにして、ミス・ハイブリダイゼ

ーションが生じ難い配列であることを確認することによって、コンピュータ等により演算処理して確認できる。このように、確認工程で合格した塩基配列を、プローブ配列として決定する。

【0022】

このように設計したプローブ配列を検出に用いるに当っては、例えば、少なくとも1本のプローブ配列に適宜の測定可能な標識物質を結合させた標識化プローブを被検サンプルと混合して、所定のハイブリダイゼーション反応を実行して、ハイブリダイズした標識物質を選択的に測定すればよい。標識物質が、FITCのような蛍光性物質であれば、適宜の蛍光測定機により、容易に測定し、データ処理手段を通じて、自動的に定量または定性分析できる。ハイブリダイゼーション反応は、必要ならば、PCRといった公知の核酸増幅反応および／またはLCRのような同定用反応と適宜組合せてもよい。また、検出の原理に応じて、適宜、同一のゲノムに関する複数種のプローブ配列を利用したり、固相支持体（微粒子、基板状チップ、カラム、フィルタ、試験紙、ウェル等）に固定してもよい。標識物質の有無または量によって、定性的または定量的に測定したら、測定データ（数値または画像）に基づいて、標的核酸分子の有無または反応量を判定し、報告書として印刷したり画面上に表示することによって、遺伝子解析の結果を得ることができる。

【0023】

なお、本発明は上述した実施の形態に限定されることなく、発明の要旨に基づいて、種々の変更が可能である。例えば、図1で説明した全ての工程を、全自動化してもよく、この場合には、各出現頻度に関するグラフ、ゲノム上のORFおよび各单位配列（ダブルおよび部分配列）の表現位置、T_m値および二次構造に関するデータやグラフ等の画面表示ないしプリントアウトについては、使用者の要望に応じて、適宜、省略して、決定したプローブ配列のみを表示または出力するだけでもよい。また、各演算工程、T_m値および二次構造の安定性の計算のみをコンピュータで行い、これら数値化および／またはグラフ化された表示画面に基づいて、使用者が抽出および取捨選択を含んだ各種評価を行って、抽出および取捨選択した情報をキーボードもしくはマウス等の入力手段を用いて入力する場

合には、各種演算ないし計算したデータをメモリ等に記憶する必要は、必ずしも無い。また、自動化の際には、オンラインを通じて、ホストコンピュータや各種病院、大学、検査センター等を結ぶネットワーク上で、各計算データや抽出または決定結果を電送したり、相互に情報交換するようにしてもよい。

【0024】

また、最も出現頻度が低い部分配列が、遺伝子上で常に一定位置に特異的に結合することが確認できた場合には、複数の低頻度部分配列に関する物理化学的条件の検討を省略してもよい。この場合でも、出現頻度の最も低い部分配列が複数有る場合には、これら複数の低頻度部分配列について、遺伝子上での結合特異性を検討するのが好ましい。また、 T_m 値と二次構造の安定性の検討は、逆の順番でも、同時でもよい。また、 T_m 値や自己二次元構造の安定性のように、物理化学的条件でプローブ配列の最適化を図る場合には、抽出した低頻度・部分配列の上記物理化学的条件の優先度に応じて、適宜、出現頻度の最下位のタプル以外で、かつ出現頻度の低い順に抽出したタプルも加えた複数の低頻度・タプルに関する部分配列を得てもよい。また、ゲノムの所定遺伝子領域として、ORF領域の全てを対象にして説明したが、ゲノム上の各種機能に応じた遺伝子領域上で、上述したようなプローブ配列を決定することによって、遺伝子の特定機能に対応するプローブ配列を有効に決定することもできる。

【0025】

【発明の効果】

本発明の方法によれば、常に正確なハイブリダイズ反応を示すような解析用配列を決定することができる。また、本発明によれば、解析用配列を迅速に決定できる。また、本発明によれば、比較的少ない計算量で段階的に解析用配列を決定するので、大型コンピュータを用いることなく、簡便にかつ経済的に有利な普及型のコンピュータで解析用配列を決定できる。

【図面の簡単な説明】

【図1】図1は、本発明の方法を実施したときの構成を模式的に示したフローチャートである。

【図2】図2は、単位配列としてのタプルを設定する方法を示す模式図である。

【図3】図3は、単位配列としての部分配列を設定する方法を示す模式図である。

【図4】図4は、単位配列を塩基配列別にして、領域内の出現頻度を示したグラフである。

【図5】図5は、全部分配列を塩基配列毎に、領域内の出現頻度を示したグラフである。

【図6】図6は、ハイブリダイゼーション環境における各種部分配列の状態を示す模式図である。

【符号の説明】

- 1 …ゲノム、
- 2 …タプル、
- 3、6、7、8 …部分配列、
- 4 …支持体、
- 5 …リンカー

【書類名】 図面

【図 1】

各 5-タブルの O R F (Open Reading Flame) 単位の出現頻度の計算



各遺伝子内の長さ 30 塩基の全ての部分配列の遺伝子特異性の評価



遺伝子特異的 30 塩基長配列の候補の決定



T_m が指定範囲内であるものを選択



安定 2 次構造を形成しないものを選択

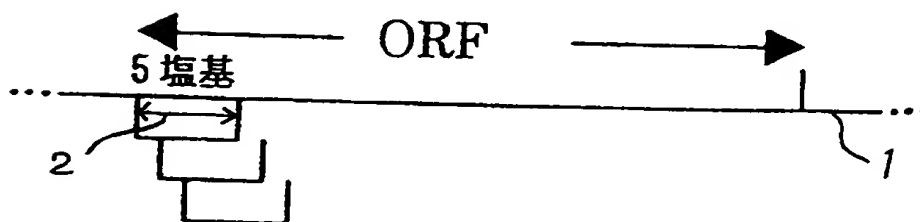


遺伝子同定プローブ結合配列の結合位置の確認

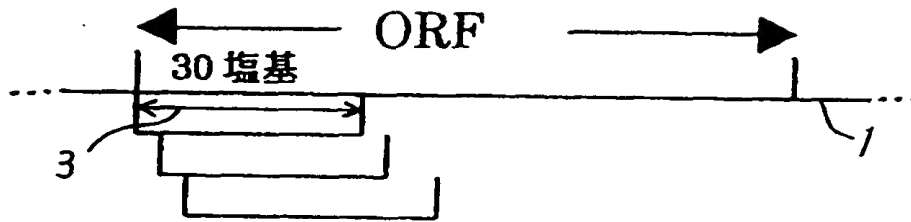


プローブ配列の決定

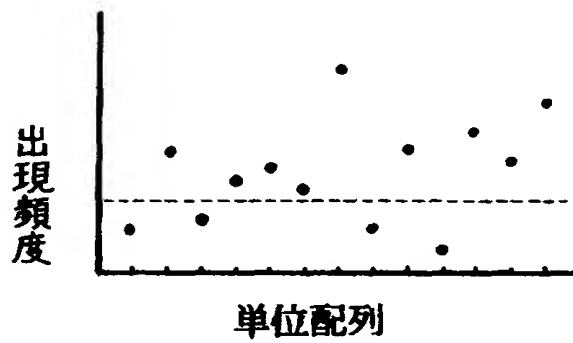
【図 2】



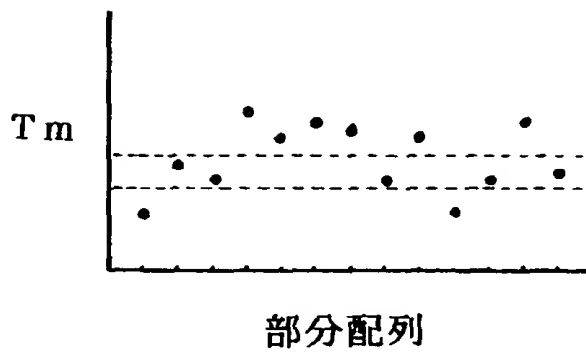
【図3】



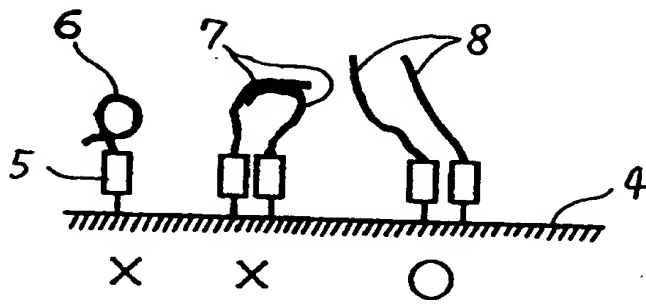
【図4】



【図5】



【图 6】



【書類名】 要約書

【要約】

【課題】常に正確なハイブリダイズ反応を示すような解析用配列を、効率良く決定する方法を提供すること。

【解決手段】限られた計算量で計算可能な5塩基毎の単位配列としてのタプル2を、ゲノム1のORF (open reading frame) 領域内に設定する。設定した全てのタプルを、塩基配列毎に出現頻度を演算して、出現頻度の低い順に抽出する。抽出した低頻度タプル2は、さらに、より長い単位配列について同様の工程を行うとともに、適宜、化学的条件による取捨選択も実行することにより、解析用配列を決定する。

【選択図】図2

出 願 人 履 歴 情 報

識別番号 [000000376]

1. 変更年月日	1990年 8月20日
[変更理由]	新規登録
住 所	東京都渋谷区幡ヶ谷2丁目43番2号
氏 名	オリンパス光学工業株式会社